In-Vehicle Coupon Recommendation System

Made By : Jay Shivankar

Github : https://github.com/jayshivankar/In-Vehicle-Coupon-Recommendation

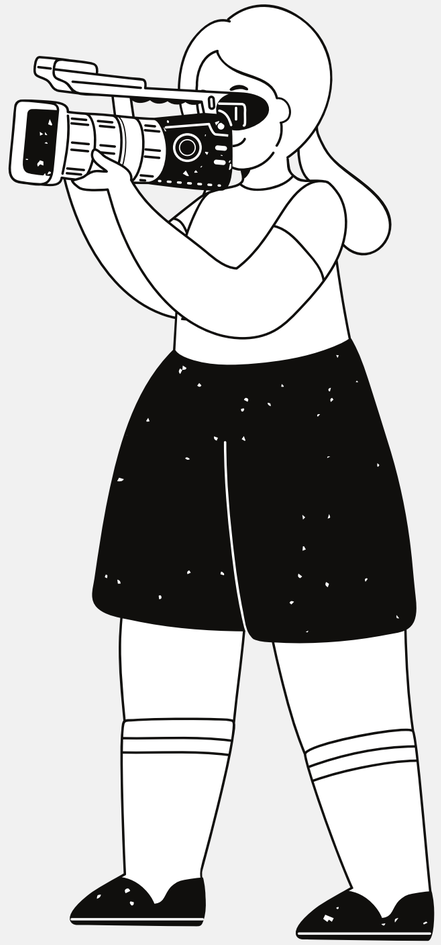
Project Overview & Methodology

Problem Statement:

- To develop a highly accurate predictive model that determines whether a driver will accept a coupon.
- The goal is to maximize marketing campaign efficiency by targeting users who are most likely to accept offers.

Dataset & Preprocessing:

- Utilized a dataset with user, context, and coupon attributes.
- Preprocessing Steps: Managed missing values, performed Label Encoding for categorical features, and used SMOTE to balance the dataset, preventing model bias.



Methodology: Model Selection

- Comprehensive Evaluation: Trained and evaluated a wide range of classification models:
- Decision Tree, Random Forest, XGBoost, AdaBoost, CatBoost, Logistic Regression, SVM, and KNN.
- Hyperparameter Tuning: Used GridSearchCV to find the optimal parameters for each model, ensuring a fair and robust comparison.
- Champion Model: After a thorough comparison, XGBoost was selected as the final model due to its superior performance across key metrics.

Exploratory Data Analysis

(Key Insights)

Overall Coupon Acceptance:

- The dataset shows a relatively balanced distribution, with approximately 57% of coupons being accepted.

What Drives Coupon Acceptance?

01. **Passenger Type:**
Drivers with friends as passengers were significantly more likely to accept coupons compared to those driving alone or with family.

02. **Coupon Type:**
Coupons for "Carry out & Take away" and "Restaurants (<$20)" had the highest acceptance rates. Bar and coffee house coupons were also popular.

03. **Occupation:**
Mostly the customers who bought the most coupons were either unemployed who had just completed graduation or were students in college or high schools.

Model Performance: Why XGBoost?

- XGBoost delivered the best overall performance after extensive hyperparameter tuning.
- It achieved the highest Test Accuracy (74.1%) and F1-Score (77.8%), indicating a superior balance of precision and recall compared to all other models, including Random Forest and CatBoost.
- XGBoost Final Performance:
- Accuracy: 0.7410
- Precision: 0.7433 (Correctly identifies coupon acceptors 74% of the time)
- Recall: 0.8169 (Finds 82% of all actual coupon acceptors)
- F1-Score: 0.7784

- Key Drivers of Prediction (Feature Importance):
- The XGBoost model identified the following as the most influential factors:
  a. expiration: How soon the coupon expires.
  b. coupon: The specific type of coupon offered.
  c. toCoupon_GEQ25min: Whether the travel time to the location is over 25 minutes.
  d. Bar: If the coupon is for a bar.

Thank you very much!